Statistics 110
B. Srinivasan
Handout #27

Autumn 2006–2007
Out: Oct. 27, 2006

Practice Midterm

Time: $\approx$ 50 minutes

Problems are arranged in roughly chronological order as well as roughly from easiest to hardest. You may leave expressions unevaluated, but you need to reduce them to the point that it would just require plugging into a calculator.

1. How many ways are there to get a straight (five cards with increasing values in a row) in a poker hand?

2. Construct a 10 point sample dataset in which the mean, median, and mode are all different.

3. A fair coin is tossed repeatedly. What is the probability that we still have not seen a head after N trials?

4. Suppose that I am waiting for a Marguerite bus whose arrival time is distributed as Exp(λ), where λ is measured in inverse minutes. At the beginning, I decide to wait a maximum of C minutes; if the bus arrives in that time then I take it and get to work in a constant time of M_1 minutes. Otherwise I walk to my destination, taking a constant time of M_2 minutes. What is my average total transit time, including both waiting and travel time?

5. Let $X \sim \text{Unif}[a, b]$ and $Y = (X - E[X])^2$. What is the variance of Y?

6. Let $P(X = \pi k) = e^{-\lambda}\frac{\lambda^k}{k!}$ for $k \geq 0$ and $\lambda > 0$. Let $Y = 1 + \sin(X)$. What is the mean of Y?

7. Suppose $F_{X,Y}(u, v) = 1 - e^{-uv}$ for $0 \leq u < \infty, 0 \leq v < \infty$. Is this a valid joint cdf? Justify your answer.

8. Assume I have a genome in which there are 30 proteins involved in respiration, 200 proteins involved in signal transduction, and 5000 proteins involved in other processes. Assay data leads me to identify a cluster of 40 genes which contain 20 proteins involved in respiration, 5 in signal transduction, and 15 from other processes. What is the probability that a cluster of genes with this composition would have been obtained via random sampling from the genome, assuming *sampling without replacement*?

9. A two pair is a poker hand which has 2 cards of the same numerical value. A royal straight flush is a poker hand which has 10/J/Q/K/A, all from the same suit. Are these two hands statistically independent? Justify your answer with a calculation.

10. Suppose you take a medical test. The test is 90% reliable in the following sense: if a person has the disease, there is a probability of .9 that the test will give a positive response. If the person does not have the disease, there is a probability of .1 that the test will give a positive response. Also assume that 1 in 100 people from this population have the disease. If you are told that a man from this population tests positive, what is the probability that he actually has the disease?